Analysis using Belief Propagation Techniques

A Report on Course Work Seminar
Submitted in partial fulfillment of the requirement of

University of Mumbai

For the Degree of

Ph.D
(Electronic Engineering)

by

Chitra Suresh

Under the guidance of

Prof.Dr.Kushal.R.Tuckley

Department of Electronics Engineering
Ramrao Adik Institute of Technology,
Sector 7, Nerul , Navi Mumbai
(Affiliated to University of Mumbai)
December 2014

Abstract

Belief Propagation Techniques uses the degree of a person's belief that an event will occur,rather than the actual probability of the event will occur. Belief Propagation Technique is used to perform inference on graphical model such as factor graphs which calculates marginal distribution for each unobserved node conditional on observed node. Belief Propagation Techniques are used for performing inference on graphical models such as Bayesian Network and Markov Random Field

Bayesian network uses the directed graphs where the directions of the arrows permits distinguish genuine dependencies for spurious dependencies induced by hypothetical observations

In Markov Random Fields,the network topology was presumed to be given and problem was to characterize the probabilistic behavior of a system complying with dependencies prescribed by network.

Belief Propagation Techniques are used in wide variety of algorithms which are used in Artificial intelligence,Signal processing and Digital communication.

Markov Random Fields and Bayesian Networks methods are used to optimize Belief Propagation algorithm.

Keywords:Belief Propagation Technique,Bayesian Network and Markov Random Field

List of Figures

Chapter 1

Introduction

This report introduces Belief Propagation Techniques and presents some of the applications.
Belief Propagation Techniques approach uses the degree of a person's belief that an event will occur,rather than the actual probability of the event will occur. Belief probabilities are properties of person's belief not the event.
A Probabilistic model is a dependency model in which relationship between each of the variable is captured.Dependency models are represent by graphical methods.
Belief Propagation Technique is used to perform inference on graphical model such as factor graphs which calculates marginal distribution for each unobserved node conditional on observed node. The graphical representation used for belief propagation Techniques are two types ,they are Markov Random Fields and Bayesian Networks.[15]
Belief Propagation Techniques are used for performing inference on graphical models such as Bayesian Network and Markov Random Field[18]
Bayesian network uses the directed graphs where the directions of the arrows permits distinguish genuine dependencies for spurious dependencies induced by hypothetical observations
In Markov Random Fields,the network topology was presumed to be given and problem was to characterize the probabilistic behavior of a system complying with dependencies prescribed by network.
Belief Propagation Techniques are used in Artificial Intelligence,Signal Processing and Digital Communication.
Some of the applications where Markov Random Fields and Bayesian Networks methods are used either as optimization tool or for implementation are discussed in chapter 3

- **Efficient Belief Propagation for Early Vision.**The method used in this application is Markov Random Fields on the Belief Propagation to solve problems for low level vision. [10]
- **Markov Network-based Unified Classifier for Face Identification** Markov Random Fields is used for face recognition for one to many identification task [5]
- **Efficient Loopy Belief Propagation using the Four Color Theorem** Four Color Theorem is based on Max-Product Belief Propagation Technique can be used in solving Markov Random Fields problems where energy is minimized. [6]

- A optimization technique which carries prioriy based message scheduling and dynamic label pruning are used for Markov Random Fields with large discrete state space is used in this application of Belief Propagation Technique **Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning** [7]

- A new stereo matching algorithm partions an image to block and optimizes with Belief propagation Technique used in this application **Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence**[2]

- A Parallel Techniques are used for implementation Belief Propagation in a acyclic factor graph are the methods used in this research **Task Parallel Implementation of Belief Propagation in Factor Graphs**[4]

- A Tile based Belief Propagation splits the Markov random field into many tiles and constructing technique is based on observation that many hypotheses is used to construct the messages are repetitive are method used in this research work **Hardware-Efficient Belief Propagation**[8]

- A method of Particle Belief Propagation (PBP) is used in **PatchMatch Belief Propagation for Correspondence Field Estimation.**[3]

- Continuous time Bayesian network method is used in **Learning continuous time Bayesian network classifiers**[9]

The organization of report is as follows Introduction in **Chapter-1**,Mathematics related to Belief Propagation Techniques is in **Chapter-2**,Different applications related to Belief Propagation Techniques are in **Chapter-3**,conclusion in **Chapter-4.**

Chapter 2

Mathematics related to Belief Propagation Technique

2.1 Probabilistic Formulation:

Bayesian Methods provide a formalism for reasoning about partial beliefs under condition of uncertainty.. In the Bayesian formalism,beliefs measures obey the three basic axioms of probability theory.

$$0 \leq P(A) \leq 1 \tag{2.1}$$

$$P(SurePropositions) = 1 \tag{2.2}$$

$$P(AorB) = P(A) + P(B) \tag{2.3}$$

if A and B are mutually exclusive
The basic expressions in the Bayesian formalism are statement about conditional probabilities be that the probability of any event A computed by conditioning it on any set of exhaustive and mutually exclusive events $B_i, i = 1, 2, ...n$

$$P(A) = \sum\{P(A \setminus B_i) \times P(B_i)\} \tag{2.4}$$

This decomposition provides the basics for the hypothetical or assumption based reasoning in the bayesian formalism.It states that the belief in any event A is weighted sum over the beliefs in all the distinct ways that event A might be realized.
Another useful generalization of the product rule is so called **Chain Rule Formula**.

$$P(A, B) = P(A|B)P(B) \tag{2.5}$$

It states that if a set of n events $(E_1, E_2, ...E_n)$

then,the probability of joint event $(E_1, E_2, ...E_n)$ can be written as a product of n conditional probabilities.

$$P(E_1, E_2, ...E_n) = P(E_n|E_{n-1}, ...E_2, E_1)......P(E_2|E_1)P(E_1) \tag{2.6}$$

This product can be derived by repeated application of equation 2.5. in any convenient order.

2.2 Probability Model

A probability model is an encoding of the probabilistic information in some domain of interest,such that each well formed sentence,statement or proposition can be calculated in accordance with axioms of probability theory.
These sentences are represented by interrelating random or stochastic variables with in model.
The probability model (M) is defined by the joint probability of all its variables or universe

Each variable in a probabilistic model may take any one of a set of mutually exclusive or collectively exhaustive states. The random variable may be discrete or continuous.The discrete random variable has finite of possible states,the probability distribution of Discrete random variable is probability mass function whereas continuous random variable takes their state from an infinite set of possible values.The probability distribution of continuous random variable is probability density function.

2.3 Probability Language

Probability language is well suited to reasoning under uncertainty.It provide a suitable frame work for processing the uncertainty relationship between variables of a probabilistic model. probability language provide consistency of inferred results and knowledge base through their axiomatic basis and provide a convenient mechanism for presenting uncertain results. The basic four primitives of probability language are

1. Likelihood
2. Conditioning
3. Relevance
4. Causation

- A likelihood of event is measure of how likely or probable it is that the event will occur ie.chance of occurrence.
- A conditioning :An event is conditional second event ,if its is changed by the knowledge of second event states.
- Relevance: events A and B are said to be relevant to each other in the context of event C ,if adding the knowledge of C to the knowledge of B,changes the likelihood of A. Example: Two events are relevance,when common sequences is observed.ie A and B are relevant if both are causes of C.
- Causation conveys a pattern of dependence between events.

2.4 Graphical Representation

A probabilistic model is dependency model in which the relationship between each of the variable is captured. A graph is denoted by G(V,E) is set of nodes or vertices V connected by the set of arcs or edge E. Graphs may be used to represent dependency model

For example :Dependency Model M comprising the variables U represented by graph. The set of arc E represents conditional dependencies between these variables.
ie.Joint probability function of M is encoded in E

Nodes of the graph corresponds to variables in dependency model U$\rightarrow$ variables

Graphical representation for probabilistic model are two types one is undirected graph and directed graph. Undirected graph had no explicit direction,an arc of influence between connected nodes Examples of undirected graph is Markov Random Fields. In directed graph arcs are either unidirectional or bidirectional directed arc provide the mechanism for representing causation Example for Directed graph is Bayesian network

Markov Random Fields and Bayesian network are two types of graphical representation for probabilistic models

2.5 Bayesian Belief Network

The Bayesian Belief Network(BBN)determines the state probabilities of each node or variable from predetermined conditional and prior probabilities.
The Direct Acyclic Graph(DAG) G(U,E)of the probability model M represents probability distribution P(U) ,where U represents set of all variables in M
Let X_1,X_2,........X_n are Variables in probability distribution P(U) Direct Acyclic Graph(DAG) in which minimal set of variables are designed as parent of each variable is X_i such that

$$P(X_i/W); W \in X_1, X_2.....X_{i-1} \tag{2.7}$$

The above equation is BBN of that probability distribution For DAG to be Bayesian network of M,it is necessary and sufficient that each variable be conditional independent of all of its non descendents given in parents also satisfies this condition.
In practice usually understands the constrains in the domain of interest.However easily identify the variables that directly influence other variables ie.eaiser to understand the local interaction of variables than the domain as a whole.
The basic concept in the Bayesian treatment of uncertainties in causal network is conditional probability ie. Given the event B,the probability of event A is x

$$P(a/b) = x \tag{2.8}$$

Conditional probability

$$P(a/b)P(b) = P(a, b) \tag{2.9}$$

In real world all events are conditioned by some context C=c

$$P(a/b, c)P(b/c) = P(a, b/c) \tag{2.10}$$

The Baye's rule also conditioned on

$$P(b/a, c) = P(a/b, c)P(b/c)/P(a/c) \tag{2.11}$$

- P(b/a,c) is posterior probability of b
- P(a/b,c) liklihood probability

- P(a/c) Normalized factor
- P(b/c) Priori probability of b

Normalised factor is not directly availableis replaced by

$$\int_b P(a/b,c)p(b/c)db \tag{2.12}$$

$$\sum_b P(a/b,c)p(b/c) \tag{2.13}$$

For continuous and discrete distributions

2.6 Markov Random Field

Markov Random Field of Belief propagation Technique is used to convey information for useful decisions and for inference problems. Markov Field of Belief propagation Technique can be expressed in terms of conditional independence of its non neighbors,once values of neighbors are known. If assigning of weights to the links of the graph must be handled with caution.If the weights are to be used translating evidential data in to meaning full probabilistic inferences such probabilistic model is both consistent and complete.Consistency guarantees that,it don't over load the graph with too many parameters.

The theory of Markov fields provides a safe method for constructing a complete and consistent quantitative model while preserving the dependency structure for an arbitrary graph (G).
The Markov Random Field method of Belief Propagation Technique consists of four conceptual steps to find marginal distribution for any probability model.

1. For unidirected graph G ,nodes are always adjust to each other.
2. For each group of nodes known as associate ,assign a nonnegative compatibility function $g_i(c_i)$,which measures the relative degree of compatibility associated with each value assignment c_i to the variable.
3. Form the product $\prod_i g_i(c_i)$ of the compatibility functions over all associates.
4. Normalize the product of all possible value combinations of the variables of the system

   $$P(x_i, ..., x_n) = K \prod_i g_i(c_i) \tag{2.14}$$

   where

   $$K = 1 \div [\sum_{x_1,...x_n} \prod_i g_i(c_i)] \tag{2.15}$$

   The normalized product P in equation 2.15 is a joint distribution that consists of all conditional independencies of graph G.

Chapter 3

Applications of Belief Propagation Techniques

Belief Propagation Technique is used for performing inference on graphical models such as Bayesian Network and Markov Random Fields.
Belief Propagation Technique calculates marginal distribution for each unobserved node, conditioning on any unobserved node.
Some of the applications where Belief Propagation Techniques are used as optimization tool are discussed in the following sections

3.1 Efficient Belief Propagation for Early Vision

Markov random field models provide a robust and unified framework for early vision problems such as stereo and image restoration. Inference algorithms based on graph cuts and belief propagation have been found to yield accurate results, but despite recent advances are often too slow for practical use.
In this paper some of the techniques are used that substantially improve the running time of the loopy belief propagation .

1. One of the techniques reduces the complexity of the inference algorithm to be linear rather than quadratic in the number of possible labels for each pixel, which is important for problems such as image restoration that have a large label set.

2. second technique speeds up and reduces the memory requirements of belief propagation on grid graphs.

3. A third technique is a multi-grid method that makes it possible to obtain good results with a small fixed number of message passing iterations, independent of the size of the input images.

Taken together these techniques speed up the standard algorithm by several orders of magnitude. In practice results obtained are as accurate as those of other global methods (e.g., using the Middlebury stereo benchmark) while being nearly as fast as purely local methods.
The three techniques are used for speeding up the belief propagation by using Markov random fields for solving low level vision problems.

The time necessary to compute a single message update from O (k^2) to O (k), where k is the number of possible labels for each pixel for the max-product formulation ie.linear rather than quadratic is possible.
By using max-product formulation and by grid graphs , The fast message updates to arbitrary discontinuity cost functions based on difference between labels and labels are embedded in some space but do not lie on a regularly spaced grid can be enhanced .

3.2 Efficient Loopy Belief Propagation using the Four Color Theorem

Recent work on early vision such as image segmentation, image denoising, stereo matching, and optical flow uses Markov Random Fields.
Although this formulation yields an NP-hard energy minimization problem, good heuristics have been developed based on graph cuts and belief propagation.
Nevertheless both approaches still require tens of seconds to solve stereo problems on recent PCs. Such running times are impractical for optical flow and many image segmentation and denoising problems and review on recent techniques for speeding them up.

Moreover in this research paper it shows that how to reduce the computational complexity of belief propagation by applying the Four Color Theorem (FCT) to limit the maximum number of labels in the underlying image segmentation to at most four. This provides substantial speed improvements for large inputs, and this for a variety of vision problems, while maintaining competitive result quality.
Statement of Four Color Theorem (FCT) is that for any 2D image there is a four-color covering regions sharing a common boundary (with more than a single point) do not have the same color, then consequence of FCT theorem is that when an image, seen as a planar graph, is segmented into contiguous regions, there are only four colors to be assigned to each pixel/node for all segments to be surrounded only by segments of different colors .
In this research paper two methods are developed based on graph cuts and belief propagation .

- In the case of Belief Propagation (BP), a key reason for its slow performance is that the algorithm complexity is proportional to both the number of pixels in the image, and the number of labels in the underlying image segmentation which is typically high. If limit the number of labels, its speed performance should improve greatly.
- By modifying the propagation algorithms like can using a low number of placeholder labels, that can reuse for non-adjacent segments. These placeholder labels can then be replaced by the full set of actual labels.
- Since image segments form a planar graph, they therefore require at most four placeholder labels by virtue of the Four Color Theorem (FCT) to still have different colors for all adjacent segments.
- A joint optimization process provides a fast segmentation through the placeholder labels and a fine grained labeling through the actual labels.

- The computational time is basically dependent on the number of placeholder rather than actual labels.

Four-Color Theorem (FCT)based on the max-product belief propagation technique can be used in early computer vision for solving MRF problems where an energy is to be minimized.
The Methods used in this research yield results that are comparable with other methods, but improve either the speed for large images and/or large label sets (the case of image segmentation, stereo matching and optical flow), or both the performance and speed (the case of image denoising)
The Four Color Theorem principle is difficult to apply in cases where the label set is discrete in the case for stereo matching and optical flow, where the disparity cost function takes discrete, unrelated values.This causes slower convergence, but proposed methods can solve the above mentioned problems also.

3.3 Markov Network-based Unified Classifier for Face Identification

In this research paper ,a novel unifying framework using a Markov network to learn the relationship between multiple classifiers in face recognition is explored.
A several complementary classifiers and assign observation nodes to the features of a query image and hidden nodes to the features of gallery images. each hidden node is connected to its corresponding observation node and to the hidden nodes of other neighboring classifiers.
For each observation-hidden node pair, to collect a set of gallery candidates that are most similar to the observation instance, and the relationship between the hidden nodes is captured in terms of the similarity matrix between the collected gallery images. Posterior probabilities in the hidden nodes are computed by the belief-propagation algorithm.
The novelty of the proposed framework is the method that takes into account the classifier dependency using the results of each neighboring classifier. The two different evaluation protocols, known and unknown image variation tests, using three different databases, which shows that the proposed framework always leads to good accuracy in face recognition.

When a face image is used as as a query, can retrieve several desired face images from a large image database,then calculate many similarities of the query image and the gallery images in the database, and the retrieved gallery images are ranked by similar orders.
It is a one-to-many identification problem and has many applications such as searching similar face images in a database and face tagging in images and videos.
Recent successful face recognition methods have attempted to merge several classifiers using multiple feature sets of different characteristics, as in component-based methods, which extract features from separate spatial regions , and heterogeneous feature-based methods which merge different domain features. These methods used the classifiers not only based on the different feature sets but also trained independently, and the similarity scores are merged with the predefined parameters. The parameter comes from the training database and it is not the best choice when the input image has different conditions. Note

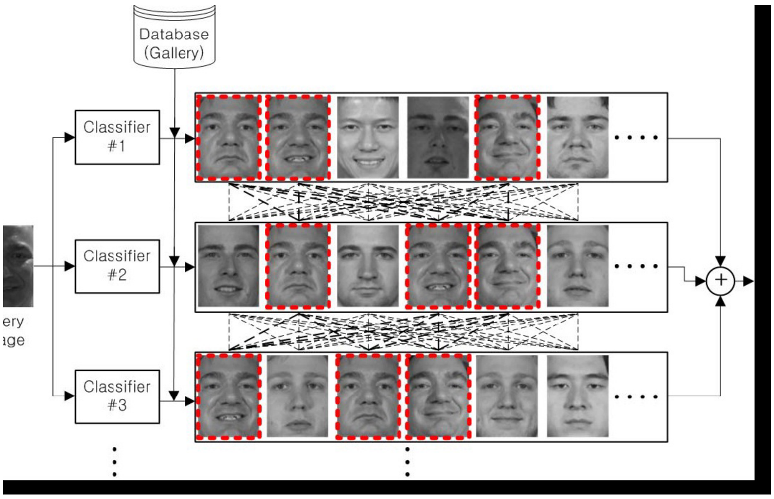

Figure 3.1: One-to-Many identification

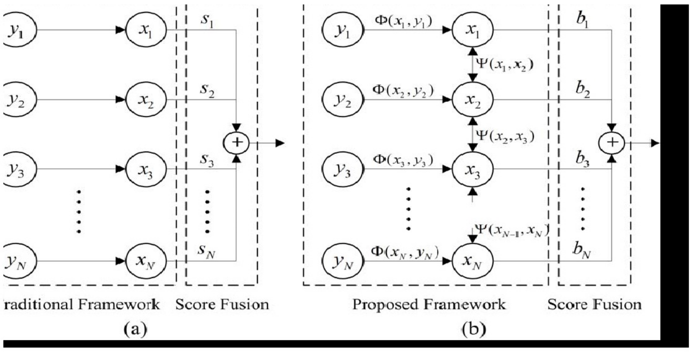

Figure 3.2: Traditional and proposed Frame Work

that these methods lead to good accuracy in face verification, but there is no specific framework for the one-to-many identification problem.

- In this paper, a novel recognition framework for the one-to-many identification issue is designed, and the simple concept is illustrated in Figure 3.1.
- First, assume that multiple classifiers have complementary characteristics, unify the multiple classifiers based not on the predefined weight values but on a Markov network, as summarized in Figure 3.2

For this purpose, assign one node of a Markov network to each classifier.The steps to find marginal probability by using Markov Fields for classifiers are as follows

1. Nodes are connected by lines, which represents the statistical dependencies
2. For an observation node, we extract a feature from a query image using the corresponding classifier.
3. At its paired hidden node, the first retrieve n similar gallery samples from the database, and their orders are made by the similarity scores for the query face image.

4. The multiple classifiers have their own lists of retrieved gallery images, which are not identical in general, thereby complementing the neighbor classifiers. Because the hidden nodes are connected by the network lines,

5. The relationship of the connected nodes is learned by the similarity scores between the neighbor classifiers, and the scores are calculated by concatenating the two gallery features of the neighbor classifiers.

6. The posterior probability at each hidden node is easily computed by the belief-propagation algorithm. Finally, marginal probability for a score value at each classifier is calculated

And also analyze the generalizability of the method using different multiple classifiers such as the Random Sampled Gabor (RSG) method, which consists of the simple and weak classifiers, and the Extended Curvature Gabor (ECG) method which consists of more complex and stronger classifiers.
The results obtained are for face recognition particularly for the one-to-many identification task, based on multiple classifiers gallery connected by a Markov network.The Markov network probabilistically models the relationships between a query and images and between neighboring gallery images. The observation-hidden node pair retrieve the similar gallery images from the database image.
The similarities between the retrieved gallery images gives the statistical dependency between the hidden nodes. Hence results obtained can be viewed as clustering-based face recognition.

3.4 Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

The problem of image completion can be defined as for a given an image which is incomplete, i.e., it has missing regions (as shown in Fig.3.1), try to fill its missing parts in such a way that a visually plausible outcome is obtained at the end. Although stating the image completion problem is very simple, the task of actually trying to successfully solve it, is far from being a trivial thing to achieve. Ideally, any algorithm that is designed to solve the image completion problem should have the following characteristics:

1. it should be able to successfully complete complex natural images

2. it should also be able to handle incomplete images with (possibly) large missing parts

3. all these should take place in a fully automatic manner, i.e., without intervention from the user.

Also, ideally any image completion algorithm to be able to handle the related problem of texture synthesis, as well. For any given a small texture as input, then asked to generate an arbitrarily large output texture, which maintains the visual characteristics of the input as shown in Fig. 3.1. It is exactly due to all of the above requirements that image completion is, in general, a very challenging problem. Nevertheless, it can be very

Figure 3.3: Object removal

Object removal is just one of the many cases where image completion needs to be applied. In the specific example shown, the user wants to remove a person from the input image on the left. He, therefore, simply marks a region around that person and that region must then be filled automatically so that a visually plausible outcome is obtained.

Figure 3.4: Image Completion

useful in many areas, e.g., it can be important for computer graphics applications, image editing, film postproduction, image restoration, etc. It has, thus, attracted a considerable amount of research over the last years. There have been three main approaches so far, for dealing with the image completion problem as shown in Fig. 3.2

1. Statistical-based methods
2. PDE-based methods
3. Exemplar-based methods

Exemplar-Based Methods : Exemplar-Based techniques, which actually have been the most successful techniques up to now. These methods try to fill the unknown region simply by copying content from the observed part of the image. All exemplar-based techniques for texture synthesis that have appeared until now, were either pixel-based or patch-based, meaning that the final texture was synthesized one pixel, or one patch at a time (by simply copying pixels or patches from the observed image, respectively).
A new exemplar-based framework which treats image completion, texture synthesis, and image inpainting in a unified manner. All image-editing tasks in the form of a discrete

global optimization is used to avoid visually inconsistent results.
The objective function of this problem is global optimization well-defined, and corresponds to the energy of a discrete Markov random field (MRF) For efficiently optimizing this MRF, a novel optimization scheme, called priority belief propagation (BP) is used which carries two very important extensions over the standard BP algorithm

1. **Priority-based message scheduling**
2. **Dynamic label pruning**

These two extensions work in cooperation to deal with the intolerable computational cost of BP, which is caused by the huge number of labels associated with MRF. Moreover, both of extensions are generic, since they do not rely on the use of domain-specific prior knowledge.They can, therefore, be applied to any MRF, i.e., to a very wide class of problems in image processing and computer vision.
As one of major limitation of the BP algorithm is its inefficiency in handling MRFs with very large discrete state spaces is considered to resolve by these extension techniques.

A novel optimization scheme which carries priority-based message scheduling and dynamic label pruning extensions for Belief Propagation known a priority BP is used.

Priority Belief Propagation can be used for other types of completion problems such as video completion or geometric completion,constrained texture synthesis
Priority-BP algorithm, which is a generic MRF optimization scheme can be used to other labeling problems for which the large cardinality of their state-space causes them to have a very high computational cost.

3.5 Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

Stereo correspondence is used in computer vision to find the depth among the cameras and objects.
This depth inference problem could be further transformed to a disparity inference problem by assuming that the cameras and objects are under epipolar geometry. The inferred disparity information could be widely applied to tracking, surveillance system, and multiview video coding
The stereo matching algorithms can be roughly divided into two categories.

- **Local approaches**
- **Global approaches**

Comparison between local approaches and global approaches

- Local approaches select disparities of image pixels using the information in a window. Therefore local approaches are faster than global approaches. However, it results in poor accuracy since the local approaches could not deal with textureless regions and occluded regions well due to the insufficient information in window.
- On the other hand, global approaches can handle the textureless and occluded regions well by formulating disparity inference as an energy minimization problem

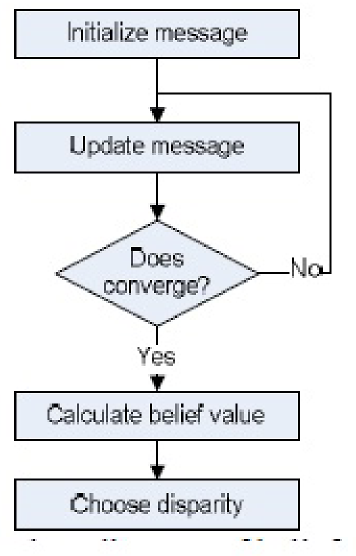

Figure 3.5: Flow diagram of Belief Propagation

Energy function and it's significance

- The energy function usually has a smoothness constraint which represents a certain physical relationship between neighboring pixel pair.
- This smoothness constraint often enforces penalty on the energy function, if the labels (disparities or segments) of neighboring pixels are inconsistent.
- Among the global methods, 2-D optimization algorithms such as graph cut and belief propagation (BP) have been applied quite successfully to optimize energy function.

The flow chart for Belief Propagation is shown in figure 3.5 The BP algorithms construct 2-D graph structures with nodes representing all the pixels in the disparity images to find the disparity map with energy closer to the global minima. However, the vast number of nodes in the 2-D graph result in extremely high computation complexity, thereby rendering 2-D optimization is too difficult to be directly implemented for real-time application.
A block-based BP algorithm that directly partitions an image into separated independent blocks. Thus, can reduce the memory size significantly due to block based computation. In addition, the independent blocks also enable parallel computation by multiple computation units. Moreover earlier convergence for each block can also improve the long running time.

A new stereo matching algorithm partitions an image to block and optimizes with belief propagation technique.This method reduces memory storage size by 99 with good performance. To enhance the interaction between neighboring blocks such that the independent block could extract useful information from neighboring finished processing blocks are possible with block belief propagation.

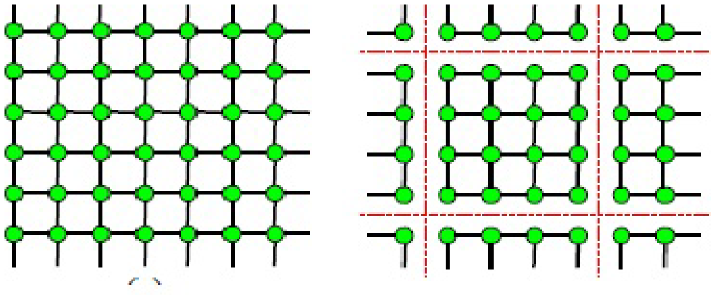

Figure 3.6: Graph of typical and block based BP
2D Graph Model is shown in figure 3.6

3.6 Task Parallel Implementation Of Belief Propagation In Factor Graphs

Graphical models have been essential tools for probabilistic reasoning. Factor graphs have emerged as a unified model of directed graphs (e.g. Bayesian networks) and undirected graphs (e.g. Markov networks).
A factor graph naturally represents a joint probability distribution that is written as a product of factors, each involving a subset of random variables. Factor graphs have found applications in Image processing,Bioinformatics and Error-control decoding used in digital communications.
Relation between Factor graphs and Belief Propagation.

- Inference is a problem of computing posterior probability distribution of certain variables given some value as observed or evidence variables.
- In factor graphs, inference proceeds with the well-known belief propagation algorithm . Belief propagation is a process of passing messages along the edges of a graph. Processing each message requires a set of operations with respect to the probability distribution of the random variables in a graph.
- Such distribution is represented by potential tables. The complexity of belief propagation increases dramatically as the number of states of variables and node degrees of a graph increase.

In many applications, such as digital communications, belief propagation must be performed in real time.
Therefore, parallel techniques are needed to accelerate the inference.Many parallel techniques have been proposed for belief propagation in factor graphs.
Parallelizing belief propagation in acyclic factor graphs still remains a challenging problem due to the precedence constraints among the nodes in the graphs.
Task scheduling is used in parallel computing is an efficient tool for linear algebra problem on general-purpose multi-core processors. The methods used for implementation of Belief Propagation using factor graphs are task dependency graph is used by using a dynamic task scheduler.

3.7 Hardware-Efficient Belief Propagation

The success of BP is due to its regularity and simplicity. It uses a simple message update process to iteratively refine the beliefs of labels for each node. A message sent from one node to another is updated according to neighboring messages and local energy functions, using simple arithmetic operations.
Loopy belief propagation (BP) is an effective solution for assigning labels to the nodes of a graphical model such as the Markov random field (MRF). The loopy BP has been widely applied to

- Stereo matching
- Image denoising
- Image inpainting

But it requires **high memory, bandwidth, and computational costs**.
However, BP algorithms generally require a great amount of memory for storing the messages, typically on the order of tens to hundreds times larger than the input data. Besides, since each message is processed hundreds of times, the saving/loading of messages consumes considerable bandwidth.
Therefore, although BP may work on high-end platforms such as desktops, it cannot be applied to most consumer electronic devices that have limited memory, computational power, and energy. It is difficult to utilize hardware parallelism to accelerate BP
The two techniques are used for sequential procedure to accelerate BP

- The first one is tile-based BP splits the Markov random field (MRF) into many tiles and only stores the messages across the neighboring tiles. The memory and bandwidth required by this technique is only a fraction of the ordinary BP algorithms. But the quality of the results comparable to other efficient algorithms ,results are tested by the publicly available Middlebury MRF benchmarks
- Second technique is that The fast message construction technique is based on the observation that many hypotheses used to construct the mesasages are repetitive. therefore, they only need to be computed once. This observation allows us to reduce the complexity of message construction from

Moreover, unlike previous sequential algorithms, the proposed algorithm can be easily parallelized. These techniques can be realized in both hardware and software. A software reference implementation compatible to the Middlebury MRF library is available online, while two hardware the first one is a very large scale integration (VLSI) circuit and the second one a graphic processing unit (GPU) program are analyzed in this paper. The techniques used to develop a tile-based message passing and fast message construction algorithm greatly reduced the memory, bandwidth, and computational costs of BP and enabled the parallel processing. With these two techniques BP can be more suitable for low-cost and power limited consumer electronics.

3.8 PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation

Patch Match is a simple, yet very powerful and successful method for optimizing Continuous labelling problems. The two main approaches used are

- The update of the solution space by sampling
- The use of the spatial neighborhood to propagate samples.

These approaches are related to steps in a specific form of belief Propagation in the continuous space, called Particle Belief Propagation (PBP). However, BP has thus far been too slow to allow complex state spaces.
The two approaches used in this research yields a new algorithm called Patch Match Belief Propagation for Correspondence Field estimation (PMBP), which is more accurate than Patch Match and orders of magnitude faster than PBP. The methods used in research is novel realistic pair wise terms that provide Smoothness used for the recent Patch Match Stereo work. The link between the popular PatchMatch method and the very well-known Belief propagation algorithm.

The link between the popular PatchMatch method and the very well-known Belief propagation algorithm introducing additional pairwise terms, These approaches can be used as both in terms of applications, such as optical flow, as well as algorithms such as different forms of message passing e.g. Treereweighted

3.9 Learning continuous time Bayesian network classifiers

Streaming data are relevant to use in finance,computer science,engineering while they are becoming increasingly important to medicine and biology.

The approach used for continuous time Bayesian network classifiers

- Continuous time Bayesian network classifiers are designed for analyzing multivariate streaming data when time duration of event matters.
- Structural and parametric learning for the class of continuous time Bayesian network classifiers are considered in the case where complete data is available.
- Conditional log-likelihood scoring is developed for structural learning on continuous time Bayesian network classifiers.
- Results show that conditional log-likelihood scoring combined with Bayesian parameter estimation outperforms marginal log-likelihood scoring.
- Conditional log-likelihood scoring becomes even more effective when the amount of available data is limited.

Conditional log-likelihood scoring function is used to learn continuous time Bayesian network classifiers from multivariate streaming data. Same function can be used for classifying multivariate trajectories in the case where the class is static .The quality of the classification performances also suggests to extend the Continuous Time Bayesian Network Classifiers to the clustering problem.

Chapter 4

Conclusion

The graphical representation or models for multidimentional probability distributions such as Markov Random Fields and Bayesian Networks are used for Belief Propagation Techniques. Some of the applications results are obtained based on following Techniques: Optimizing Markov Random Fields to speed up the Belief Propagation used for face recognition and low level vision problem

Another optimizing scheme which carries priority based message scheduling and dynamic label pruning extensions are used to handle Markov Random Field with large discrete space.

The block based Belief Propagation algorithm reduces memory size significantly due to block based computation.

A parallel techniques are used for implementation of Belief Propagation Techniques in acyclic factor graphs.

A tile based Belief Propagation and fast construction techniques are realized both in Software and Hardware.

Learning continuous time Bayesian network classifiers are used for data streaming by using conditional log-likelihood scoring function.

Bibliography

[1] Judea Pearl `Probabilistic Reasoning in Intelligent Systems` Kaufmann Morgann Publishers, Second Revised And updated Edition,San Francisco, California, 1998.

[2] Yu-Chang Andnelson Chang `Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence`, 1-4244-1017-7187,IEEE Transisition

[3] F.Besse,C.Rother,A.Fitzgibbon,J.Kautz `Pmbp:Patchmatch Belief Propagation For Correspondence Field Estimation` University College Of London U.K.,Microsoft Research Cambridge ,2012

[4] Nam Ma,Yinglong Xia,Viktorprasanna `Task Parallel Implementation Of Belief Propagation In Factor Graphs` This Research Was Suppored By U.S.National Foundation Under Grant No.Cns-101880

[5] Wonjun Hwang Kyungshik,Roh Junmo Kim `Markov Network-Based Unified Classifier For Face Identification` This ICCV 2013 Paper is Open Access Available In IEEE Xplore

[6] Radu Timofte,Lucvan Gool `Efficient Loopy Belief Propagation Using The Four Color Theorem` Isics,Esat-Psi/Iminds V,Belgium,Computer Vision Lab,Zurich,Switzerland

[7] Nikos Komadakis,Georgios Tziritas `Image Completion Using Efficient Belief Propagation Via Priority Scheduling And Dynamic Pruning` IEEE Transactions In Image Processing ,Volume;16,No.11,November 2007

[8] Chia-Kai Liang,Chao-Chung Cheng,Yen-Cheieh Lai `Hardware-Efficient Belief Propagation` IEEE Transactions On Circuits And Systems For Video Technology Vol;21,No.5,May 2011

[9] Daniele Codecasa,Fabiostella `Learning Continuous Time Bayesian Network Classifiers` International Journal Of Approximate Reasoning,55(2014) 1728-1746

[10] Pedro F.Felzenszwalb,Daniel P.Huttenlocher `Efficient Belief Propagation For Early Vision` International Journal Of Computer Vision (Springer Science)70(1),41-54,2006

[11] Dan Yuan `Understanding Belief Propagation And Its Applicatons` Department Of Computer Engineering,University Of California

[12] Erik.B.Sudderth And William T.Freeman `Signal And Image Processing With Belief Propagation`

Accepted To Appear In IEEE Signal Processing Magazine

[13] N.Kmodakis And G.Tziritas `Image Completion Using Efficient Bp Via Priority Scheduling And Dynamic Pruning` IEEE Trans.In Image Process,Vo.16 No.11

[14] A,Brunton.C.Shu And G.Roth `Belief Propagation On Gpu For Streo Vision` IEEE Proceedings Crv 2006.

[15] Hans Andrea Loeliger `An Introduction To Factor Graphs` IEEE Signal Processing Magazine,2004

[16] Jianguo Ding `Probabilistic Inferences In Bayesian Networks` International Journal Of Computer Science,Arxiv:1011.0935V2,5 November 2010

[17] Frank R.Kschischang,Brendan J.Frey,Hans Andrea Loeliger `Factor Graph and The Sum Product Algorithm` IEEE Transcations On Information Theory,Vol47.No.2,Feb,2001

[18] Judea Pearl `Reverend Bayes On Inference Engines:A Distributed Hierarchical Approach` Aaai-B2 Proceedings 1982

[19] Heckerman.D. `A Tutorial On Learning With Bayesian Networks` Technical Report Mst-Tr-95-96,Microsoft Corporation ,Redmond,Wa